WHAT IS CLAIMED IS:

1. A computer-implemented method for enriching sparse data for machine learning, comprising:

receiving the sparse data;

enriching the received data around a deviation of the mean of the received data using a predetermined distribution; and

outputting the enriched data for unbiased learning and improved performance during the machine learning.

2. The method of claim 1, wherein machine learning comprises:

supervised artificial neural network learning.

3. The method of claim 1, further comprising:

checking the received data for sparseness; and

enriching the checked data around the deviation of the mean of the received data based on the outcome of the checking.

4. The method of claim 1, wherein checking the received data further comprises:

comparing the received data with a predetermined number.

5. The method of claim 4, wherein enriching the received data further comprises:

enriching the received data around the deviation of the mean of the received data based on the outcome of the comparison.

6. The method of claim 1, further comprising:

rearranging the received data based on class.

7. The method of claim 6, further comprising:

normalizing the rearranged data based on attributes in the rearranged data.

8. The method of claim 6, further comprising:

checking each class of data in the rearranged data for sparseness; and

enriching each class of data around a deviation of the mean associated with the respective class based on the outcome of the checking.

9. The method of claim 8, wherein checking each class of data further comprises:

comparing each class of data to a predetermined number.

10. The method of claim 9, wherein enriching each class of data comprises:

enriching each class around a deviation of the mean associated with the respective class based on the outcome of the comparison.

11. The method of claim 10, wherein enriching each class around a deviation of the mean associated with the respective class further comprises:

computing the mean and standard deviation for each class of data in the rearranged data; and

generating additional data for each class using the associated computed mean and standard deviation.

12. The method of claim 11, wherein generating additional data further comprises:

generating additional data between limits computed using the equation:

$$\bar{x} \pm k\sigma$$

wherein $\bar{x}$ is the computed mean associated with each class, k is a constant varying between 0.25 to 3, and σ is the computed standard deviation associated with each class.

13. The method of claim 12, wherein the predetermined distribution further comprises:

arranging the enriched data using the equation:

$[X_{mn}]\ [W] = [B_i]$

wherein W is a weight matrix, X is input patterns, and B_i's are the classes; and

rearranging in the max-min-max pattern:

Let for class ***i***

$(Rx_{1N} - Rx_{2N}) > (Rx_{2N} - Rx_{3N}) > \ldots > (Rx_{(i-1)N} - Rx_{iN}) =$

$(Rx_{(i+2)N} - Rx_{(i+1)N}) < (Rx_{(i+3)} - Rx_{(i+2)}) < \ldots < (Rx_{AN} - Rx_{(A-1)N})$

where $Rx_{1N} \rightarrow$ *Row* x_{1N} are enriched data values.

14. The method of claim 1, wherein the predetermined distribution comprises distributions selected from the group consisting of normal distribution, exponential distribution, logarithmic distribution, chi-square distribution, t-distribution, and F-distribution.

15. The method of claim 1, wherein the received data comprises data selected from the group consisting of static data and real-time data.

16. The method of claim 15, further comprising:

if the received data is static data, then reading a sample of the received static data using a predetermined window length; and

if the received data is real-time data, then reading a sample of the received real-time data using a dynamically varying window of predetermined window length.

17. The method of claim 16, further comprising:

if the received data is real-time data, then repeating the reading of the sample of the received real-time data using a dynamically varying window of predetermined window length.

18. A computer readable medium having computer-executable instructions for performing a method of machine learning when only sparse data is available, comprising:

enriching the sparse data around a deviation of the mean of the received data using a predetermined distribution; and

outputting the enriched data for unbiased machine learning.

19. The computer readable medium of claim 18, wherein machine learning comprises:

supervised artificial neural network learning.

20. The computer readable medium of claim 18, further comprising:

checking the received data for sparseness; and

enriching the received data around the deviation of the mean of the received data based on the outcome of the checking.

21. The computer readable medium of claim 18, wherein checking the received data further comprises:

comparing the received data with a predetermined number.

22. The computer readable medium of claim 21, wherein enriching the received data further comprises:

enriching the received data around the deviation of the mean of the received data based on the outcome of the comparison.

23. The computer readable medium of claim 18, further comprising:

rearranging the received data based on class.

24. The computer readable medium of claim 23, further comprising:

normalizing the rearranged data based on attributes in the rearranged data.

25. The computer readable medium of claim 23, further comprising:

checking each class of data in the rearranged data for sparseness; and

enriching each class of data around a deviation of the mean associated with the respective class based on the outcome of the checking.

26. The computer readable medium of claim 25, wherein checking the each class of data further comprises:

comparing each class of data to a predetermined number.

27. The computer readable medium of claim 9, wherein enriching the each class of data comprises:

enriching each class around a deviation of the mean associated with the respective class based on the outcome of the comparison.

28. The computer readable medium of claim 27, wherein enriching each class around a deviation of the mean associated with the respective class further comprises:

computing the mean and standard deviation for each class of data in the rearranged data; and

generating additional data for each class using the associated computed mean and standard deviation.

29. The computer readable medium of claim 28, wherein generating additional data further comprises:

generating additional data between limits computed using the equation:

$$\bar{x} \pm k\sigma$$

wherein $\bar{x}$ is the mean associated with each class, k is a constant varying between 0.25 to 3, and σ is the standard deviation associated with each class.

30. The computer readable medium of claim 29, wherein the predetermined distribution further comprises:

arranging the enriched data using the equation:

$[X_{mn}]\ [W] = [B_i]$

wherein W is a weight matrix, X is input patterns, and B_i's are the classes; and

rearranging in the max-min-max pattern:

Let for class ***i***

$(Rx_{1N} - Rx_{2N}) > (Rx_{2N} - Rx_{3N}) > \ldots > (Rx_{(i-1)N} - Rx_{iN}) =$

$(Rx_{(i+2)N} - Rx_{(i+1)N}) < (Rx_{(i+3)} - Rx_{(i+2)}) < \ldots < (Rx_{AN} - Rx_{(A-1)N})$

where $Rx_{1N} \rightarrow Row\ x_{1N}$.

wherein Rx_{1N} is the first row of the *Xth (reference)* class consisting of N features, Rx_{2N} is the second row of the *Xth* (reference) class consisting of N features, and so on.

31. The method of claim 18, wherein the predetermined distribution comprises distributions selected from the group consisting of normal distribution, exponential distribution, and logarithmic distribution.

32. The computer readable medium of claim 18, wherein the received data comprises data selected from the group consisting of static data and real-time data.

33. The computer readable medium of claim 32, further comprising:

if the received data is static data, then reading a sample of the received static data using a predetermined window length ; and

if the received data is real-time data, then reading a sample of the received real-time data using a dynamically varying window of predetermined window length.

34. The computer readable medium of claim 33, further comprising:

if the received data is real-time data, then repeating the reading of the sample of the received real-time data using a dynamically varying window of predetermined window length.

35. A computer system for a machine learning in a sparse data environment, comprising:

a storage device;

an output device; and

a processor programmed to repeatedly perform a method, comprising:

receiving the data;

enriching the received data around a deviation of mean of the received data using a predetermined distribution; and

outputting the enriched data for unbiased machine learning.

36. The system of claim 35, wherein machine learning comprises:

supervised artificial neural network learning.

37. The system of claim 35, further comprising:

rearranging the received data based on class.

38. The system of claim 37, further comprising:

normalizing the rearranged data based on attributes in the rearranged data.

39. The system of claim 37, further comprising:

checking each class of data in the rearranged data for sparseness; and

enriching each class of data around a deviation of the mean associated with the respective class based on the outcome of the checking.

40. The system of claim 39, wherein checking the each class of data further comprises:

comparing each class of data to a predetermined number.

41. The system of claim 40, wherein enriching each class of data comprises:

enriching each class around a deviation of mean associated with the respective class based on the outcome of the comparison.

42. The system of claim 41, wherein enriching the each class around a deviation of mean associated with the respective class further comprises:

computing the mean and standard deviation for each class of data in the rearranged data; and

generating additional data for each class using the associated computed mean and standard deviation.

43. The system of claim 42, wherein generating additional data further comprises:

generating additional data between limits computed using the equation:

$$\bar{x} \pm k\sigma$$

wherein $\bar{x}$ is the mean associated with each class, k is a constant varying between 0.25 to 3, and σ is the standard deviation associated with each class.

44. The system of claim 35, wherein the predetermined distribution comprises distributions selected from the group consisting of normal distribution, exponential distribution, and logarithmic distribution.

45. A computer-implemented system for machine learning in a sparse data environment, comprising:

a receive module to receive sparse data;

an analyzer to enrich the received data around a deviation of the received data using a predetermined distribution; and

an output module coupled to the analyzer to output the enriched data for unbiased learning and increased performance during machine learning.

46. The system of claim 45, further comprising:

a database coupled to the receive module to receive and store sparse data.

47. The system of claim 45, wherein machine learning comprises:

supervised artificial neural network learning.

48. The system of claim 45, further comprising:

a comparator coupled to the analyzer to check the received data for sparseness, wherein the analyzer enriches the checked data around the deviation of the mean of the received data based on the outcome of the checking.

49. The system of claim 48, wherein the comparator checks the received data for sparseness by comparing the received data with a predetermined number.

50. The system of claim 49, wherein the analyzer enriches the received data around the deviation of the mean of the received data based on the outcome of the comparison.

51. The system of claim 50, wherein the analyzer rearranges the received data based on class.

52 The system of claim 51, wherein the analyzer normalizes the rearranged data based on attributes in the data.

53. The system of claim 51, wherein the analyzer checks each class of data for sparseness, and enriches each class around a deviation of the mean associated with the class based on the outcome of the checking by the analyzer.

54. The system of claim 53, wherein the comparator compares each class in the rearranged data with a predetermined number and wherein the analyzer enriches each class around a deviation of the mean associated with the class based on the outcome of the comparison by the comparator.

55. The system of claim 54, wherein the analyzer enriches data in each class by computing a mean and standard deviation for each class in the rearranged data, and the analyzer further generates additional data for each class based on the respective computed mean and standard deviation.

56. The system of claim 55, wherein the analyzer generates additional data between limits computed using the equation:

$$\bar{x} \pm k\sigma$$

wherein $\bar{x}$ is the mean associated with each class in the rearranged data, k is a constant varying between 0.25 to 3, and σ is the standard deviation associated with each class in the rearranged data.

57. The system of claim 56, wherein the analyzer further computes additional data using the equation:

$[X_{mn}]\ [W] = [B_i]$

wherein W is a weight matrix, X is input patterns, and B_i's are the classes; and rearranging in the max-min-max pattern:

Let for class ***i***

$(Rx_{1N} - Rx_{2N}) > (Rx_{2N} - Rx_{3N}) > \ldots > (Rx_{(i-1)N} - Rx_{iN}) =$

$(Rx_{(i+2)N} - Rx_{(i+1)N}) < (Rx_{(i+3)} - Rx_{(i+2)}) < \ldots < (Rx_{AN} - Rx_{(A-1)N})$

where $Rx_{1N} \rightarrow Row\ x_{1N}$.

58. The system of claim 45, wherein the received data comprises data selected from the group consisting of static data and real-time data.

59. The system of claim 58, further comprising:

a reading module coupled to the receive module reads a sample of the received data having a predetermined window length.

60. The system of claim 59, wherein the reading module reads the sample of the received data using a predetermined window length when the read data is static data, and reads a sample of the received data using a dynamically varying window of predetermined window length when the read data is real-time data.

61. The system of claim 60, wherein the reading module repeats the reading of the sample of the received data using a dynamically varying window of predetermined window length when the received data is real-time data.

62. The system of claim 45, further comprising:

a unique numeric transformation module coupled to the database to extract words from text stored in the database and to transform each of the extracted words into a unique numerical representation.